



L'évaluation interactive des apprentissages dans l'enseignement supérieur : intérêt pédagogique et limites

Véronique Bedin, Dominique Broussal

► To cite this version:

Véronique Bedin, Dominique Broussal. L'évaluation interactive des apprentissages dans l'enseignement supérieur : intérêt pédagogique et limites. Les Dossiers des sciences de l'éducation , 2011, 26, pp.133-144. hal-00762102

HAL Id: hal-00762102

<https://hal.science/hal-00762102>

Submitted on 6 Dec 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

L'évaluation interactive des apprentissages dans l'enseignement supérieur : intérêt pédagogique et limites

Véronique Bedin¹
Dominique Broussal²
Université de Toulouse, UT2, CREFI-T
France

Résumé

L'université française voit d'un œil dubitatif l'introduction de pratiques évaluatives, qui s'inscrivent pourtant dans une perspective affichée d'amélioration de la qualité. Les réserves exprimées à l'encontre de l'évaluation des enseignements par les étudiants, rejoignent la frilosité de la pédagogie universitaire en matière de développement de nouvelles pratiques évaluatives, relativement aux apprentissages des étudiants. Le dispositif ici étudié, consiste en une évaluation interactive des apprentissages d'étudiants de master de sciences de l'éducation. Celui-ci combine une évaluation mutuelle des étudiants entre eux, et une phase de co-évaluation au cours de laquelle les points de vue des étudiants sont confrontés à ceux des enseignants-chercheurs. A rebours d'une conception parfois idéalisée de l'évaluation formative, il s'agit, dans cette recherche, d'identifier précisément quels sont les effets formatifs du dispositif, mais aussi de revenir sur un certain nombre de limites. La visée est tout à la fois de contribuer au développement d'une pédagogie universitaire et d'interroger plus spécifiquement les enjeux liés à l'interactivité en situation d'évaluation d'apprenants adultes.

■ **Mots clés : pédagogie universitaire, master de sciences de l'éducation, évaluation interactive, évaluation mutuelle, co-évaluation, apprentissages**

Introduction

L'enjeu de cet article est de montrer, dans un contexte universitaire dans lequel l'évaluation occupe une place limitée et peu reconnue, en quoi et comment des pratiques d'évaluation des apprentissages, qualifiées d'« interactives », peuvent justement participer de la formation d'étudiants d'un master de sciences de l'éducation. Comment de telles pratiques développent-elles, notamment, les compétences évaluatives de ces étudiants, tant dans leurs relations avec leurs pairs qu'avec les enseignants-chercheurs ? La comparaison entre les effets escomptés du dispositif évaluatif proposé et ceux finalement observés, conduira à interroger sa valeur formative et apprenante, plaçant ainsi les finalités et les modalités de la pédagogie universitaire et de ses traductions évaluatives au coeur du débat. Le niveau master pris en

¹ Maître de conférences en sciences de l'éducation à l'Université de Toulouse II-Le Mirail (UT2), Centre de recherches en éducation, formation et insertion de Toulouse (CREFI-T).

² Id.

considération complexifie l'analyse. Il se prête mal à des applications *stricto sensu* des théories de l'apprentissage, telles qu'elles ont été élaborées en milieu scolaire ou en formation continue. Ce sont les résultats d'une recherche qualitative menée en 2010 à l'Université de Toulouse-Le Mirail, auprès d'une promotion d'étudiants de ce second cycle, qui seront ici exposés.

1. L'évaluation encore « en chantier » dans l'université française

Comme d'autres systèmes de formation et contextes socio-professionnels, l'université est soumise à l'exigence d'évaluer la qualité de ses activités et performances. La construction d'un espace européen d'enseignement supérieur renforce cette perspective. Dire que la « culture » de l'évaluation, récemment introduite dans les universités européennes, n'a pas encore trouvé ses marques dans tous les établissements relève toutefois d'un euphémisme (Bedin, 2009). Des formes intégrées d'évaluation telles que les conseillait déjà le Comité national d'évaluation (Friedberg & Musselin, 1992) avaient connu des difficultés de mise en oeuvre. Que dire alors des évaluations actuelles de l'Agence d'évaluation de la recherche et de l'enseignement supérieur, qui sont encore plus centralisées et externalisées ?

Ce constat déficitaire vaut pour différents types d'évaluation, qu'il s'agisse de l'évaluation des enseignements par les étudiants (EEE), de l'évaluation des acquis des étudiants ou de modalités évaluatives universitaires autres d'ailleurs, comme nous le verrons dans cette contribution.

Concernant l'évaluation des enseignements par les étudiants, Dejean (2006, p. 61) souligne « les réticences à l'évaluation de l'enseignement en France », malgré l'arrêté du 9 avril 1997 qui prévoyait pourtant sa mise en œuvre dans les établissements d'enseignement supérieur. La culture professionnelle des enseignants-chercheurs marquée par la « liberté académique » constituant un obstacle, selon l'auteur. Des résistances qui peuvent se transformer en positionnement critique clairement assumé, en invoquant « une idéologie de l'évaluation » pensée comme une « grande imposture » (Zarka, 2009, p. 3).

Quant à l'évaluation des acquis des étudiants, Romainville (2002, 2006) tire les conclusions suivantes : si ces pratiques y sont à la fois quantitativement conséquentes, trop diversifiées et sélectives pour les étudiants, elles n'en demeurent pas moins paradoxalement peu connues, non explicitées et d'une qualité relative. Au total, elles restent plus régies par des « allants de soi » et des caractéristiques disciplinaires que par une réelle professionnalité et le souci de l'employabilité des diplômés dans le choix des compétences à évaluer.

La recherche évaluative n'échappe pas non plus à un bilan mitigé : dans l'enseignement supérieur, elle est cantonnée à un champ « travaillé à la marge ou de façon dispersée » (Fave-Bonnet, 2010, p. 8).

L'expression de « chantier » que nous avons utilisée dans le titre du paragraphe illustre la représentation que l'on peut posséder de la situation actuelle : l'omnipotence symbolique et latente de l'évaluation à l'université et le contexte incertain dans lequel elle s'actualise, un défi que des enseignants-chercheurs de sciences de l'éducation auront disciplinairement à relever. Parmi les questions vives qui se posent au plan pédagogique, une retient plus particulièrement notre attention dans cet article : comment l'évaluation peut-elle apporter une valeur ajoutée au processus d'apprentissage des étudiants de master ? Elle relève de l'évaluation des apprentissages des étudiants lors de leurs enseignements.

2. Le « pari » formatif de l'évaluation interactive dans la pédagogie universitaire

Dans cette mise en œuvre déjà difficile des pratiques évaluatives (EEE, évaluation des apprentissages des étudiants notamment), celles relevant de la pédagogie universitaire restent les moins investies par les enseignants-chercheurs, prisonniers, selon Romainville (1996, p. 151), du « couple maudit de l'université », l'enseignement et la recherche. Les conclusions des analyses convergent quant à l'évaluation de leurs activités professionnelles (Langevin & Bruneau, 2000 ; Dejean, 2002 ; Musselin, 2008) : la priorité est accordée aux activités scientifiques sur celles de l'enseignement, les premières sont jugées plus stratégiques et porteuses en termes de reconnaissance professionnelle et de promotion, ce qui correspond effectivement à la réalité constatée. Là également, la situation française se démarque par rapport à d'autres pays francophones comme le Canada, la Suisse ou la Belgique par exemple, dans lesquels des services universitaires proposent du conseil pédagogique aux enseignants du supérieur. Tout en soulignant le caractère émergent de ces activités, Rege Colet (2006, p. 185) montre qu'elles « interviennent en véritable soutien à l'enseignement et se positionnent comme facilitatrices de la transformation de l'enseignement universitaire et de l'innovation pédagogique ».

Malgré un contexte globalement peu incitateur, comme nous l'avons indiqué, c'est pourtant un des objets de la pédagogie universitaire que nous avons choisi d'étudier, en mettant l'accent sur la place que peut occuper l'évaluation dans le processus d'apprentissage des étudiants, dans des conditions, bien entendu, à expliciter. L'évaluation est posée, dans ce cas de figure, « comme une pratique pédagogique et didactique » (Talbot, 2009, p. 85) et répond alors à des finalités formatives sur lesquelles nous reviendrons *infra*. Les étudiants auxquels nous ferons référence sont ceux d'un master de sciences de l'éducation. La spécificité de la formation à et pour la recherche de ce type de diplôme, questionne sans concession et de manière inédite les pratiques pédagogiques. Comment faire acquérir ou renforcer des méta-compétences (analyse et vigilance critiques, conceptualisation, auto et co-réflexivité, évaluation formative) ou des postures caractéristiques de l'identité d'un étudiant de deuxième cycle (Broussal & Chabridon, 2009), dans son rapport aux autres acteurs universitaires et à lui-même ?

La « didactisation » de la situation pédagogique par la mise en œuvre d'un dispositif socio-cognitif complexe d'évaluation interactive, qui sera détaillé dans le paragraphe suivant, devait permettre de répondre à cet objectif ambitieux. Il s'agit d'une évaluation doublement interactive qui mobilise un outil d'évaluation chaque fois, entre étudiants d'abord, entre étudiants et enseignants-chercheurs ensuite. Cet outil est considéré comme un « référentiel » au sens où l'entend Figari (1994, p. 44) : « un système de références » qui n'est pas nécessairement fixé une fois pour toutes, « un appareil construit » et à reconstruire. Le terme « interactif » renvoie aux travaux d'Allal (2007, p. 13), qui reprennent d'ailleurs des recherches antérieures, et qui mettent en exergue les interactions de l'apprenant avec les acteurs de son environnement et/ou avec des outils qui favorisent des démarches autorégulatrices. Pour ce qui est de notre propos, cette évaluation interactive recouvre, dans un premier temps, une évaluation dans laquelle les étudiants évaluent, à partir d'un outil, les travaux de leurs pairs ; elle est, dans ce cas, qualifiée d'« évaluation mutuelle », en ce sens où « deux ou plusieurs sujets évaluent réciproquement les démarches, produits, projets, etc., formulés par chacun d'entre eux » (Allal & Michel, 1993, cité par Paquay, Darras & Saussey, 2001, p. 122). Dans un second temps, cette évaluation est confrontée à celle effectuée par les enseignants-chercheurs sur le même produit et correspond alors à une « co-évaluation », définie comme un « dialogue qui se noue entre l'autoévaluation réalisée par le sujet et

l'évaluation que le formateur a effectuée à propos du même objet » (*ibid.*, p. 122). Le troisième temps est consacré à l'attribution d'une note finale et individuelle qui valide l'unité d'enseignement concernée, tout en tenant compte des résultats du processus évaluatif qui précède.

Cette démarche évaluative, en accordant un statut prépondérant aux interactions sociocognitives et à la stimulation intellectuelle par le truchement d'outils, s'inscrit dans le paradigme du socioconstructivisme, et revêt, à ce titre, une fonction formative, sur laquelle il convient néanmoins de s'interroger. De nombreux auteurs, spécialistes de l'évaluation et cités en bibliographie, ont spécifié les caractéristiques de l'évaluation formative, particulièrement dans le cadre de l'enseignement du premier degré d'ailleurs. Elle est appelée « évaluation-régulation » par Bonniol et Vial (1997/2000), car elle participe effectivement de la régulation des apprentissages (développement de savoirs, de compétences et de postures), en permettant notamment à l'étudiant de recueillir de manière continuée des informations sur les processus et/ou les résultats en cours, d'en apprécier la portée et de lui permettre ainsi de s'adapter à la situation pédagogique proposée, ce qui tend à augmenter sa capacité réflexive, à le responsabiliser et à le faire progresser.

Dans le cas qui nous intéresse, le dispositif étudié est censé favoriser le développement de l'apprentissage des compétences évaluatives chez les étudiants de master, ce qui sous-tend leur capacité à objectiver différents éléments constitutifs d'une situation, à argumenter un point de vue dans une perspective collaborative, à se positionner institutionnellement dans une unité d'enseignement de deuxième cycle, etc. L'enquête montrera la spécificité de la régulation produite dans le cadre d'une évaluation mutuelle ou d'une co-évaluation sur cet apprentissage, de même que l'impact cognitif de l'outil d'évaluation par rapport à des interactions de nature plus sociale.

Le terme de « pari », qui prolonge celui de « chantier » du premier paragraphe, met l'accent sur le caractère non systématique des effets positifs produits, toujours dépendants des conditions de mise en œuvre et des éléments de contexte, particulièrement en deuxième cycle universitaire, face à un groupe d'« apprenants » qui n'est pas réductible à un groupe d'élèves. L'évaluation formative court toujours le risque d'en rester à « une utopie porteuse » (Hadji, 1997). Les résultats de l'investigation empirique permettront effectivement de nuancer le « tout formatif » dont on pourrait penser qu'une évaluation interactive est « intrinsèquement » dotée.

3. La présentation du dispositif d'évaluation interactive proposé aux étudiants de master

L'enquête de terrain s'est déroulée en mai 2010 auprès de 30 étudiants inscrits dans une unité d'enseignement (UE) du master des sciences de l'éducation de l'Université de Toulouse II - Le Mirail (France), qui visait à les sensibiliser à l'analyse de situations de formation.

Au démarrage de cette UE de 25 heures en octobre 2009, deux situations de formation vécues par les étudiants ont d'abord été oralement présentées par leurs soins et débattues ensuite en grand et en petits groupes pour donner enfin lieu à la production de dossiers collectivement réalisés et soutenus devant l'ensemble de la promotion. Chaque dossier proposait une analyse critique et argumentée d'une des situations de formation exposées.

En ce qui concerne l'évaluation, les étudiants devaient, par petits groupes, évaluer la soutenance orale des dossiers de leurs pairs, à l'aide d'un outil d'évaluation multi-critères

élaboré par les enseignants-chercheurs. L'enjeu était moins d'évaluer la soutenance des dossiers en elle-même que de placer les étudiants en situation d'« exercice » évaluatif et de les confronter alors à la réalité et aux difficultés d'une expérience d'évaluation interactive, de manière à pouvoir l'analyser ensuite. La majorité d'entre eux n'avait jamais rencontré une situation « pédagogique » de ce type, conçue ici comme un dispositif incitatif au développement de compétences évaluatives et d'analyse critique, dispositif faisant office de « prétexte » à évaluer d'une certaine manière. Les étudiants avaient à attribuer une note de 0 à 10 par critère, ainsi qu'une note globale (évaluation mutuelle et *a priori* formative). Cette activité d'une vingtaine de minutes avait lieu après chaque soutenance de dossier. Sept dossiers ont été présentés dans le cadre d'une seule journée. L'outil, portant le titre de référentiel d'évaluation, comprenait cinq critères : 1) le contenu de l'exposé, 2) l'élocution, 3) la posture et l'attitude générale, 4) la syntaxe et le vocabulaire, 5) le respect du temps de l'exposé. L'attribution d'une note, dans le cas présent, relevait plus « d'une détermination d'une valeur » qu'une application, au sens strict, du modèle quantitatif de la mesure, modèle que Lecoine (1997, p. 23) a complexifié en montrant comment « la démarche même de mesure » participait « de l'établissement et de l'évolution de ce qui est reconnu comme valeur de référence » (*ibid.*, p. 25). La notation des étudiants était ensuite confrontée à celle des enseignants-chercheurs responsables de l'unité d'enseignement qui avaient utilisé le même outil pour l'effectuer (co-évaluation et évaluation *a priori* formative). L'ensemble du processus se soldait, *in fine*, par l'obtention ou non de l'unité concernée (évaluation sommative et certificative).

« L'exercice » évaluatif proposé aux étudiants répondait à une visée formative, entendue au sens de mise en oeuvre d'une activité régulatrice, entre pairs d'abord, et entre étudiants et enseignants-chercheurs ensuite, en vue d'apprécier collectivement la qualité de la soutenance des dossiers et sur chacun des critères retenus du référentiel d'évaluation. Activité réflexive qui permettait effectivement de repérer les « erreurs » commises lors de la soutenance et les améliorations à y apporter, de confronter des argumentations, d'objectiver et d'ajuster des points de vue jusqu'à la prise de décision concertée de l'attribution de notes. Cette situation est considérée comme apprenante du point de vue de l'apprentissage de compétences évaluatives par les étudiants, dans un modèle de « l'évaluation-régulation » comme nous avons pu antérieurement le définir. L'opportunité offerte aux formés de s'exercer à porter un jugement sur la valeur de la soutenance des dossiers, de délibérer jusqu'à établir le choix le plus raisonné possible justifiant la notation, a constitué la valeur formative du dispositif proposé. Le contexte institutionnel d'évaluation, réel et non simulé, a renforcé, de surcroît, la densité symbolique de la situation pédagogique choisie.

Le recueil de données s'est effectué *a posteriori*, quelques mois après l'expérimentation décrite *supra*. Il a pris la forme d'une grille d'analyse que les étudiants ont librement complétée. Le terme de « grille d'analyse » renvoyait à la finalité auto-réflexive et d'analyse critique visée par l'outil mis en place qui ne se réduisait donc pas à un simple recueil de données. La grille d'analyse constituait un élément à part entière de l'ensemble de la démarche évaluative proposée. Il était ainsi demandé aux étudiants, individuellement et par écrit, de donner leur avis sur le dispositif d'évaluation auquel ils avaient antérieurement participé : l'intérêt perçu, les difficultés rencontrées, les apprentissages réalisés, la qualité de l'outil d'évaluation multi-critères utilisé et son mode d'usage, l'appréciation du rôle des pairs tout au long de la démarche, les modes de travail collaboratifs mis en oeuvre, le rapport aux enseignants-chercheurs et à leur propre mode d'évaluation, l'appréciation des valeurs en jeu dans l'« exercice » évaluatif effectué et les améliorations à y apporter.

Dix-neuf étudiants seulement sur trente inscrits à cet enseignement ont restitué le matériau d'enquête. Sept documents sur dix-neuf n'étaient pas suffisamment complétés pour

être correctement traités, notamment les réponses aux questions ouvertes. Ce sont donc douze grilles qui ont finalement été retenues, de manière à constituer un corpus que nous qualifierons scientifiquement « d'exploitable ». L'investigation empirique choisie s'inscrit dans une méthodologie de type qualitatif, qui met l'accent sur un choix limité mais pertinent de sujets, sur l'analyse thématique comme moyen de traitement des données textuelles recueillies et sur une approche compréhensive du processus évaluatif étudié. Dans cette configuration méthodologique, priorité est accordée à la perception qu'ont les étudiants de la valeur formative du dispositif d'évaluation proposé et des apprentissages que ce dispositif leur a permis de développer, dans tous les cas de figure « expérimentés » (évaluation mutuelle, co-évaluation, évaluation relayée par un outil pour l'essentiel).

4. Des apprentissages envisagés aux apprentissages perçus : quel constat, quelles perspectives d'amélioration ?

Ainsi que nous l'avons évoqué, une rubrique de l'enquête interrogeait explicitement les étudiants sur les apprentissages réalisés au cours des activités d'évaluation mutuelle et de co-évaluation. L'analyse des résultats a conduit à dresser un premier inventaire de ces apports :

- apprendre à noter constitue pour eux le principal intérêt formatif du dispositif, compétence qu'ils appréhendent sous l'angle exclusivement métrologique. Un tel constat appelle un commentaire. L'exigence d'avoir à produire une note à l'issue de l'évaluation mutuelle, créait certes les conditions d'un tel investissement. Le fait que chacun des critères donne lui-même lieu à une note peut être perçu comme un renforcement de cette orientation métrologique explicite. Il aurait été possible de neutraliser cette dimension métrologique, en proposant aux étudiants de formuler des appréciations qualitatives, dans le but de favoriser l'émergence de pratiques évaluatives relevant d'autres modèles (valeur, reconnaissance). Or un choix différent a été fait. Il consistait à parier sur la conjugaison potentielle des différents paradigmes évaluatifs, au motif que c'est dans cette intrication que les compétences évaluatives s'exercent, dans de nombreux contextes professionnels. Cette vision a été abordée lors des enseignements, mais la réflexion des étudiants n'en porte pas trace. Le retour à une conception concurrentielle ou pour le moins dissociée des différents modèles, s'explique selon nous par la difficulté réelle qu'il y a à se maintenir dans des postures intellectuelles affirmant la coexistence des contraires ou la conflictualité des visées, face à l'exigence unificatrice de la pratique.
- laisse supposer de la part des étudiants une vision concurrentielle des différents modèles, aux dépens de conceptions qui défendent une approche complémentaire : parce qu'elle renvoie au modèle de la mesure, l'activité de notation pourrait avoir ainsi vidé l'activité évaluative de ses autres potentialités.
- le travail avec les pairs, lors de l'évaluation mutuelle, a permis à de nombreux étudiants de découvrir l'hétérogénéité des conceptions et des perceptions de chacun. Il les a conduit à s'interroger sur les mécanismes de négociation inhérents à la prise de décision collective ;

- le mode d'évaluation expérimenté est envisagé, par certains, comme un moyen de participer aux changements de pratiques pédagogiques au sein de l'université, en faisant notamment « émerger de nouvelles formalités d'évaluation ».

En confrontant ces résultats avec les intentions pédagogiques qui avaient présidé à la mise en place du dispositif, dans un master orienté vers la recherche dans les champs de l'éducation et de la formation, nous avons pu identifier un certain nombre de limites imputables au dispositif lui-même ou à ses modalités de mise en œuvre. Nous les présenterons dans les paragraphes qui suivent.

Il nous a, par ailleurs, semblé utile d'accorder un traitement spécifique à des points de vue qui, s'exprimant dans l'enquête, remettent en cause de façon directe et parfois brutale l'intérêt même de l'activité proposée, ou la qualité de l'accompagnement réalisé par les enseignants-chercheurs auprès du groupe. Ces critiques nous intriguent et nous intéressent en effet, tout autant par leur nature que par leur caractère imprévisible. Rien dans les échanges qui ont eu lieu à l'occasion des moments formels d'évaluation interactive, n'avaient pu nous les laisser prévoir. Ce qui s'exprime à la faveur du dispositif de recherche réinterroge le dispositif évaluatif dans son principe comme dans sa contractualisation tacite. Cela nous conduira à faire un pas de côté, et à considérer sous un jour inattendu le rôle que la consultation *a posteriori* des étudiants a pu jouer. Nous l'envisagerons non plus dans sa fonction compréhensive ou heuristique, mais comme participant de façon périphérique au projet évaluatif.

5. Le référentiel à l'épreuve de la pratique d'évaluation mutuelle : la référencement comme perspective

Munis de la grille d'évaluation multi-critères (contenu, élocution, posture et attitude générale, syntaxe et vocabulaire, respect du temps de l'exposé) qui leur avait été communiquée par les enseignants-chercheurs, les étudiants de master se sont donc trouvés en situation de noter collectivement la prestation orale de leurs pairs. Signalons qu'aucune réticence ou inquiétude ne s'étaient exprimées lors de la présentation préalable de la grille et des consignes de travail, eu égard par exemple à la faisabilité de l'exercice. L'enquête nous révèle pourtant que seuls cinq étudiants sur douze avaient eu dans le passé une pratique d'évaluation, les autres en ayant seulement fait l'expérience, pour reprendre la distinction opérée par Vial (2009, p. 12). Il est dès lors instructif de voir que, malgré sa brièveté, l'activité évaluative à laquelle ils ont pris part, a suffi pour qu'ils remettent en cause l'idée d'une pratique ordinaire, aisément accessible. Les difficultés éprouvées ont fait émerger le constat selon lequel une modification de posture est à accomplir, « *allant de l'évalué à l'évaluateur* ». Plusieurs insistent après coup sur l'importance de former l'évaluateur, exprimant qu'ils ne sont « *pas assez armés pour donner une note* ».

Ce que montre en outre, de façon extrêmement marquée, l'investigation empirique, c'est que la mise au travail des critères proposés, dans le cadre d'une pratique finalisée, a conduit à discuter leur caractère opérationnel. Cette dénaturalisation du référentiel a initié une réflexion portant sur le processus même de « référentialisation » (Figari, 1994), c'est-à-dire sur la modélisation de l'évaluation par la mise en lien des référents, des critères et des indicateurs. Ainsi, un étudiant indique avoir appris qu'il est utile de « *définir et construire l'outil pour que l'objectivité soit la plus présente, fiable et claire* » ; un autre reconnaît qu'en dépit du cadre commun fourni par la grille, « *chaque individu peut avoir ses propres critères* ». D'autres encore se projettent dans l'élaboration de « *critères plus précis et*

d'avantage orientés sur la qualité du travail en groupe des étudiants », ou se questionnent « sur la pertinence des items d'évaluation ». La production du référentiel émerge donc comme une dimension stratégique de la pratique évaluative requérant des compétences spécifiques.

Cette phase de dénaturalisation du référentiel, telle qu'elle se réalise à travers la mise à l'épreuve d'une grille élaborée par les enseignants-chercheurs, nous paraît à conserver en l'état. Dans une perspective d'amélioration du dispositif, il semblerait intéressant d'y adjoindre une deuxième phase au cours de laquelle les étudiants pourraient proposer un certain nombre de modifications portant sur la grille initiale. L'objectif n'étant pas tant l'amélioration du support de notation, que la négociation effective des critères et indicateurs.

Une deuxième proposition concerne la diversification des modèles de référence mobilisés par les étudiants et leur conscientisation. Nous l'évoquions plus haut, la mise en question du référentiel et la prise de conscience de la posture spécifique de l'évaluateur, ainsi que de son « *métier* », interviennent à l'intérieur d'un modèle dominant qui s'exprime sous diverses formes, sans pour autant être nommé en tant que tel, ni remis en cause. Les termes de « *rigueur* », de « *neutralité* », d'« *objectivité* », de « *distanciation* » ou de « *précision* » y prévalent. Il s'agit du modèle historique de la mesure, dont on conçoit qu'il ait été induit par l'activité de notation requise, mais dont on aurait pu souhaiter qu'il fût questionné dans une perspective de référencement (Vial, 2009). Ce travail de référencement se distingue de l'opération de référentialisation évoqué ci-dessus, leurs enjeux n'étant pas de même nature. La référencement se donne en effet pour objectif de mettre en lumière les références intellectuelles dans lesquelles s'inscrit l'étudiant-évaluateur et qui l'« *agissent* ». Un tel processus, parce qu'il suppose une maîtrise des connaissances théoriques en matière d'évaluation suffisante pour en dégager les logiques profondes, pour faire émerger les modèles de pensée parfois implicites, imposerait sans doute de la part des enseignants-chercheurs un étayage important. Enrichir la nature de l'activité attendue, en ajoutant à la notation d'autres formes de productions évaluatives, pourrait faciliter l'identification des *épistémès* et des visions du monde qui président à l'un ou à l'autre modèle.

6. La valeur du dispositif : les co-évaluateurs et leurs doubles

Nous disions notre étonnement d'avoir découvert, lors du dépouillement de l'enquête, les points de vue de plusieurs étudiants remettant en cause la valeur du dispositif et la qualité même de l'accompagnement proposé. Ainsi un étudiant, particulièrement explicite, indique que « *cette notation est toujours aussi inutile et aléatoire et ne sert qu'à la compétition entre les individus, à créer des relations de pouvoir entre les étudiants et à les formater* ». Un autre souligne le manque d'investissement des enseignants-chercheurs « *sur la mise en place de cette expérience de co-décision* ». D'autres sont plus allusifs mais rejoignent certains éléments de diagnostic. S'autorisant de la promesse paritaire du tiret que porte l'injonction co-évaluative, les évalués se sont fait évaluateurs. Rien que d'attendu ! Mais pris dans ce jeu de miroir, l'objet de l'évaluation s'est insidieusement troublé, puisque c'est de l'activité des enseignants-chercheurs et non plus de l'apprentissage des pairs que ces évaluateurs audacieux prétendent traiter. Faut-il pour autant rejeter ces propos ? L'anonymat garanti par l'enquête entache-t-il leur vertu régulatrice ? Ce type d'expression est-il compatible avec la dissymétrie constitutive de la relation d'enseignement que le contexte universitaire renforce encore ? L'enquête n'est-elle qu'un exutoire, une façon de s'élever contre le pouvoir sur lequel toute pratique d'évaluation institutionnelle se fonde, ou dont elle se recommande ?

De telles questions obligent à prendre la mesure des effets que génère le dispositif d'évaluation interactive, rajoutant à la complexité inhérente à l'évaluation. Ainsi la co-

existence d'une logique de contrôle visant « la transparence, la technicité, l'outillage » et d'une logique du « reste » dans laquelle prévalent « le flou, le mal discernable, l'opacité des relations humaines » (Vial, 2009, p. 70), qui est au cœur de nombre de pratiques évaluatives, prend-elle un relief particulier sous l'effet contaminant du « co- ». En installant un scénario paritaire dans lequel chacun pèserait du même poids, il révèle paradoxalement la disparité fondamentale des valences et invite à la discuter. Le « jeu » évaluatif engagé ne peut dès lors perdurer sans que se trouve explicitée la dimension instituante de la dyssymétrie. Une telle perturbation n'est pas dépourvue d'intérêt, pour des apprentis chercheurs en formation et en éducation. Elle les invite en effet à élargir le spectre de leur analyse des dispositifs pédagogiques, en les sensibilisant à des considérations stratégiques ou politiques. Une telle exploitation, par la dynamique critique qu'elle engage, nécessiterait sans doute une conception singulière de l'enseignement universitaire, plus attachée à l'*auctoritas* qu'à la *potestas*, non pas tant soucieuse d'exercer cette dernière que de « garantir aux sujets de venir à l'énonciation » (Lecoq, 2001, p. 205). Elle supposerait par ailleurs une conception élargie de la régulation des apprentissages qui, en faisant droit aux questionnements sur le sens et la valeur de l'activité, conduirait à interroger les attendus respectifs du métier d'étudiant et du métier d'enseignant-chercheur.

Conclusion

Au terme de cette étude, quels constats peut-on dresser à partir de cette expérience d'évaluation interactive ?

En accord avec les effets escomptés de la démarche socio-constructiviste, la situation-problème à laquelle les étudiants ont été confrontés, au cours de la situation d'évaluation mutuelle, les a amenés à échanger sur leurs conceptions et à les faire évoluer. La verbalisation *a posteriori* témoigne du développement de connaissances opératoires et d'une amorce d'appropriation des outils sémiotiques. Le pari formatif semble donc tenu. Il faut toutefois remarquer l'inégale mobilisation dans l'action des ressources théoriques disponibles. Le modèle métrologique l'emporte ainsi sur les autres, l'activité de notation empêchant la mise au travail des autres modèles (valeur, reconnaissance). Ce constat d'une dissociation des registres épistémiques et pragmatiques, s'il n'a rien de surprenant (Pastré, 2006), conduit toutefois à s'interroger sur les visées de l'enseignement universitaire, dans un contexte où la professionnalisation fait parfois figure d'exercice obligé. Quel rôle les modèles théoriques de l'évaluation prétendent-ils jouer par exemple sur l'évolution des pratiques évaluatives réelles et de leurs finalités ? Comment les enseignants-chercheurs explicitent-ils pour un public visant à se professionnaliser dans les champs de l'éducation ou de la formation le rapport entre la production scientifique et les contextes socio-professionnels ?

Concernant ce qui pourrait être référé à de l'ingénierie pédagogique, il apparaît, en deuxième lieu, que la recherche sur l'évaluation interactive à l'Université gagnerait à se démarquer de celle qui est conduite en contexte scolaire pour explorer, à côté des dimensions apprenantes et méta-cognitives de tels dispositifs, ce qui relève de la formation d'une praxis évaluative.

Bibliographie

- Allal, L. & Michel, Y. (1993). Autoévaluation et évaluation mutuelle en situation de production écrite. In L. Allal, D. Bain & P. Perrenoud (Ed.), *Évaluative formative et didactique du français* (pp. 7-23). Neuchâtel : Delachaux et Niestlé.
- Allal, L. (2007). Régulations des apprentissages : orientations conceptuelles pour la recherche et la pratique en éducation. In L. Allal & L. Mottier Lopez L. (Ed.), *Régulation des apprentissages en situation scolaire et en formation* (pp. 207-234). Bruxelles : De Boeck.
- Bedin, V. (2009) (Dir.). *L'évaluation à l'université. Évaluer ou conseiller ?* Rennes : Presses Universitaires de Rennes.
- Bonniol, J.-J. & Vial, M. (1997/2000). Les modèles de l'évaluation. Bruxelles : De Boeck Université.
- Broussal, D., Chabridon, M. (2009). La place des valeurs dans le développement professionnel d'étudiants de master : point de vue des stagiaires et de leurs référents. *Travail et apprentissages*, 3, 26-43.
- Dejean, J. (2002). *L'évaluation de l'enseignement dans les universités françaises*. Rapport établi à la demande du Haut conseil de l'évaluation de l'école. Paris : Haut conseil de l'évaluation de l'école. Disponible sur Internet : <http://cisad.adc.education.fr/hcee/documents/rapport5journalistes.pdf>.
- Dejean, J. (2006). Les réticences à l'évaluation de l'enseignement en France, signe de la culture professionnelle des enseignants-chercheurs ou trait de la culture française. In M. Romainville & N. Rege Colet (Ed.), *La pratique enseignante en mutation à l'université* (pp. 61-80). Bruxelles : De Boeck.
- Fave-Bonnet, M.-F. (2010) (Dir.). *L'évaluation dans l'enseignement supérieur en questions*. Paris : L'Harmattan.
- Figari, G. (1994). *Évaluer : Quel référentiel ?* Bruxelles : De Boeck Université.
- Friedberg, E. & Musselin, C. (1992). *Le gouvernement des universités. Perspectives comparatistes*. Paris : L'Harmattan.
- Hadji, C. (1997). *L'évaluation démystifiée*. Paris : ESF.
- Langevin, L. & Bruneau, M. (2000). *Enseignement supérieur. Vers un nouveau scénario*. Issy-les Moulineaux : ESF.
- Lecointe, M. (1997). *Les enjeux de l'évaluation*. Paris : L'Harmattan.
- Lecoq, D. (2001). Formation et management : quelques éléments conceptuels pour une analyse de pratiques. In C. Blanchard-Laville & D. Fablet (Ed.), *Sources théoriques et techniques de l'analyse des pratiques professionnelles* (179-207). Paris : L'Harmattan.
- Musselin, C. (2008). *Les universitaires*. Paris : Éditions La Découverte.

- Paquay, L., Darras, E. & Saussez, F. (2001). Les représentations de l'autoévaluation. In G. Figari & M. Achouche (Ed.), *L'activité évaluative réinterrogée, regards scolaires et socioprofessionnels* (pp. 119-133). Bruxelles : De Boeck.
- Pastré, P. (2006). Apprendre à faire. In E. Bourgeois & G. Chapelle Gaëtane (Ed.), *Apprendre et faire apprendre* (pp. 109-121). Paris : PUF.
- Rege Colet, N. (2006). Représentations et modèles pédagogiques des conseillers pédagogiques en milieu universitaire. In M. Romainville & N. Rege Colet (Ed.), *La pratique enseignante en mutation à l'université* (pp. 185-198). Bruxelles : De Boeck.
- Romainville, M. (1996). Enseignement et recherche, le couple maudit de l'université. *Gestion de l'enseignement supérieur*, 8, 2, 151-160.
- Romainville, M. (2002). *L'évaluation des acquis dans l'enseignement universitaire*. Rapport établi à la demande du Haut conseil de l'évaluation de l'école. Paris : Haut conseil de l'évaluation de l'école. Disponible sur Internet : <http://cisad.adc.education.fr/hcee>.
- Romainville, M. (2006). Quand la coutume tient lieu de compétence : les pratiques d'évaluation des acquis à l'université. In M. Romainville & N. Rege Colet (Ed.), *La pratique enseignante en mutation à l'université* (pp. 19-40). Bruxelles : De Boeck.
- Talbot, L. (2009). *L'évaluation formative. Comment évaluer pour remédier aux difficultés d'apprentissage ?* Paris : Armand Colin.
- Vial, M. (2009). *Se former pour évaluer*. Bruxelles : De Boeck.
- Zarka, Y.-C. (2009). L'évaluation : un pouvoir supposé savoir. *Cités*, 37, 113-123.